

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ  
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ  
імені ІГОРЯ СІКОРСЬКОГО»**

**ФАКУЛЬТЕТ ПРИКЛАДНОЇ МАТЕМАТИКИ**

Кафедра системного програмування і спеціалізованих комп'ютерних систем

«До захисту допущено»

Завідувач кафедри

\_\_\_\_\_  
(підпис)      Тарасенко В.П.  
(ініціали, прізвище)

“ \_\_\_\_ ” \_\_ червня \_\_\_\_ 2019 р.

**Дипломний проект  
на здобуття ступеня бакалавра**

з напрямку підготовки      6.050102 «Комп'ютерна інженерія»

на тему: «Алгоритм і програма розпізнавання реквізитів банківських карток».

Виконала: студентка IV курсу, групи KB-51

Костюк Ірина Сегрійвна

\_\_\_\_\_  
(підпис)

Керівник доцент, к.т.н Потапова К.Р.

\_\_\_\_\_  
(підпис)

Консультант з нормоконтролю доцент, к.т.н. Клятченко Я.М.

\_\_\_\_\_  
(підпис)

Рецензент д.т.н., проф. каф. ОТ ФІОТ Сергієнко А.М.

\_\_\_\_\_  
(підпис)

Засвідчую, що у цьому дипломному  
проекті немає запозичень з праць інших  
авторів без відповідних посилань.

Студент \_\_\_\_\_  
(підпис)

Київ – 2019 року

# **Пояснювальна записка до дипломного проекту**

на тему: «Алгоритм і програма розпізнавання реквізитів банківських карток»

Київ – 2019 року

## ВІДОМІСТЬ ДИПЛОМНОГО ПРОЕКТУ

№ з/п	Формат	Позначення	Найменування	Кількість листів	Примітка
1	A4	ІАЛЦ. 045490.000	Завдання на дипломний проект	2	
2	A4	ІАЛЦ. 045490.001 ОА	Опис альбому	2	
3	A4	ІАЛЦ. 045490.002 ТЗ	Технічне завдання	4	
4	A4	ІАЛЦ. 045490.003 ТП	Відомості технічного проекту	2	
5	A4	ІАЛЦ. 045490.004 ПЗ	Пояснювальна записка	52	
6	A4	ІАЛЦ. 045490.005 Д1	Розпізнавання реквізитів банківських карток. Схема алгоритму	1	
7	A4	ІАЛЦ. 045490.006 Д2	Класифікація даних. Схема алгоритму	1	
8	A4	ІАЛЦ. 045490.007 Д3	Тренування моделі розпізнавання. Схема алгоритму	1	
9	A4	ІАЛЦ. 045490.008 Д4	Взаємодія модулів програми. Схема структурна	1	

				ІАЛЦ.045490.000		
	ПІБ	Підп	Дата	Алгоритм і програма розпізнавання реквізитів банківських карток. <b>Відомість</b> дипломного проекту	Лист	Листів
Розробн.	Костюк І.С.				1	1
Керівн.	Потапова К.Р.				КПІ ім. Ігоря Сікорського Каф. СПіСКС Гр. KB-5x	
Консульт						
Н/контр.	Клятенко Я. М.					
Зав.каф.	Тарасенко В.П.					

## Зміст

1.	<u>НАЙМЕНУВАННЯ І ОБЛАСТЬ ЗАСТОСУВАННЯ</u>	2
2.	<u>ПІДСТАВА ДЛЯ РОЗРОБКИ</u>	2
3.	<u>МЕТА І ПРИЗНАЧЕННЯ РОБОТИ</u>	2
4.	<u>ДЖЕРЕЛА РОБОТИ</u>	2
5.	<u>ТЕХНІЧНІ ВИМОГИ</u>	3
5.1.	<u>Вимоги до системи, що розробляється</u>	3
5.2.	<u>Вимоги до апаратного забезпечення</u>	3
5.3.	<u>Вимоги до програмного забезпечення</u>	3
6.	<u>ЕТАПИ РОЗРОБКИ</u>	4

					ІАЛЦ.045490.002 ТЗ			
Змін	Арк.	№ докум.	Підпис	Дата				
Розроб.		Костюк І.С.			Алгоритм і програма розпізнавання реквізитів банківських карток.  <b>Технічне завдання</b>	Літ.	Аркуш	Аркушів
Перевір.		Потапова К.Р.					1	4
Н. контр.		Клятченко Я.М.				КПІ ім. Ігоря Сікорського, ФПМ,КВ-51		
Затвер.		Тарасенко В. П.						

## **1. НАЙМЕНУВАННЯ І ОБЛАСТЬ ЗАСТОСУВАННЯ**

Найменування роботи – «Алгоритм і програма розпізнавання реквізитів банківських карток».

Область застосування: комп'ютерний зір.

## **2. ПІДСТАВА ДЛЯ РОЗРОБКИ**

Підставою для розробки є завдання на виконання дипломного проекту першого (бакалаврського) рівня вищої освіти, затверджене кафедрою системного програмування та спеціалізованих комп'ютерних систем Національного технічного університету України «Київський Політехнічний Інститут ім. Ігоря Сікорського».

## **3. МЕТА І ПРИЗНАЧЕННЯ РОБОТИ**

Метою даного дипломного проекту є розробка алгоритму і програми для розпізнавання реквізитів банківських карток на зображеннях, що повинна максимально точно розпізнавати номер, власника і термін дії банківської карти за її зображенням.

## **4. ДЖЕРЕЛА РОБОТИ**

Джерелами інформації для розроблення є технічна література з областей комп'ютерного зору та машинного навчання, публікації у періодичних виданнях та електронні статті у мережі Інтернет, що порушують питання розробки.

					ІАЛЦ.045490.002 ТЗ	Арк.
						2
Изм.	Лист	№ докум.	Підпис	Дата		

## 5. ТЕХНІЧНІ ВИМОГИ

### 5.1. Вимоги до програми, що розробляється

Програмна система повинна забезпечувати наступні функції:

- Завантаження зображення з файлової системи;
- Отримання зображення за допомогою веб-камери;
- Розпізнавання реквізитів банківської карти;
- Виведення отриманих результатів розпізнавання на екран;

### 5.2. Вимоги до апаратного забезпечення

Комп'ютер на базі процесора сімейства AMD, A4-4020 та вище або сімейства Intel, G3900 та вище з оперативною пам'яттю 256 Мбайт і більше.

### 5.3. Вимоги до програмного забезпечення

1. Операційна система одного з сімейств:

- Windows XP або вище;
- Linux 3.0 або вище;
- Mac OS X 10.7.3 (Lion) або вище.

2. Додаткові інсталяційні пакети.

- Python3;
- OpenCV;
- Numpy;
- Tkinter.

					ІАЛЦ.045490.002 ТЗ	Арк.
						3
Изм.	Лист	№ докум.	Підпис	Дата		

## 1. ЕТАПИ РОЗРОБКИ

№ з/п	Назва етапів роботи та питань, які мають бути розроблені відповідно до завдання	Термін виконання
1.	Видача завдання на дипломне проектування	18.12.2018
2.	Розробка технічного завдання	05.02.2019
3.	Аналіз існуючих рішень	14.03.2019
4.	Вибір середовища розробки	25.03.2019
5.	Розробка програмного продукту	20.04.2019
6.	Відлагодження програмного продукту	24.04.2019
7.	Підготовка пояснювальної записки	13.05.2019
8.	Оформлення матеріалів проекту	20.05.2019
9.	Попередній огляд матеріалів диплому на кафедрі	23.05.2019

## АНОТАЦІЯ

Метою даного дипломного проекту є створення алгоритму і програми для розпізнавання реквізитів банківських карток.

У ході розробки виконаний аналіз існуючих методів та алгоритмів виявлення та розпізнавання символів на зображеннях. На базі виконаного теоретичного аналізу сформовано вимоги до програмної системи, розроблено алгоритм та програму, що дозволяє розпізнавати реквізити банківських карток за його попередньо завантаженим зображенням або моментально отриманою фотографією з веб-камери. Також був реалізований графічний інтерфейс.

Програма розроблена мовою програмування Python з використанням бібліотеки алгоритмів лінійної алгебри NumPy та бібліотеки машинного навчання і комп'ютерного зору OpenCV. Для реалізації користувацького інтерфейсу використана графічна бібліотека Tkinter.

*Ключові слова:* комп'ютерний зір, розпізнавання символів, банківська карта, Python, OpenCv, NumPy.



## **ABSTRACT**

The aim of the diploma project is creation of an algorithm and program for recognition of bank card details.

During the development, was performed the analysis of existing methods and algorithms for detecting and recognizing characters on images. On the basis of the theoretical analysis, was formed the requirements for the software system, was developed an algorithm and a program that allows to recognize bank card details from his pre-loaded image or instantly received photos from a webcam. Graphical user interface was also implemented.

The system is developed using the Python programming language and the library of algorithms of linear algebra NumPy as well as machine learning and computer vision library OpenCV. To implement the user interface, was used the Tkinter graphics library.

*Key words:* computer vision, character recognition, bank card, Python, OpenCv, NumPy.



[illegible]

## ЗМІСТ

СПИСОК ТЕРМІНІВ, СКОРОЧЕНЬ ТА ПОЗНАЧЕНЬ.....	2
ВСТУП.....	4
1. АНАЛІЗ ІСНУЮЧИХ РІШЕНЬ ТА ОБГРУНТУВАННЯ ТЕМИ ДИПЛОМНОГО ПРОЕКТУ .....	6
1.1. Розпізнавання образів.....	6
1.2. Розпізнавання тексту .....	8
1.3. Розпізнавання реквізитів банківської карти.....	10
1.4. Аналіз існуючих рішень .....	13
1.5. Обґрунтування теми дипломного проекту .....	18
2. ХАРАКТЕРИСТИКА ІНСТРУМЕНТАРІЮ ТА ОПИС ПРОЦЕСУ РОЗРОБКИ ПОГРАМИ .....	21
2.1. Методи та алгоритми попередньої обробки зображення .....	21
2.2. Штучний інтелект .....	23
2.3. Машинне навчання .....	24
2.4. Штучна нейронна мережа .....	27
2.5. Принципи та методи навчання .....	34
2.6. Обґрунтування вибраного інструментарію .....	43
3. ОПИС РОЗРОБЛЕНОЇ ПРОГРАМИ.....	45
ВИСНОВКИ.....	50
СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ .....	51

Додаток 1. Копії графічних матеріалів

Додаток 2. Лістинг програми

Додаток 3. Слайди презентації

					ІАЛЦ.045490.004 ПЗ								
Змін	Анк	№ док.ум	Підпис	Дата									
Розроб.		Костюк І.С.			Алгоритм і програма розпізнавання реквізитів банківських карток. <i>Пояснювальна записка</i>			Літ.		Аркуш		Аркушів	
Перевір.		Потапова К.Р.								1		52	
								КПІ ім. Ігоря Сікорського, ФПМ.КВ-51					
Н. контр.		Клятченко Я.М.											
Затвер.		Тарасенко В. П.											

## СПИСОК ТЕРМІНІВ, СКОРОЧЕНЬ ТА ПОЗНАЧЕНЬ

**OCR** (Optical Character recognition) – це процес перекладу графічного зображення символів в комп’ютерні текстові символи.

**БПК** – банківська платіжна картка.

**ШІ** - штучний інтелект.

**ML** (Machine Learning) – підрозділ штучного інтелекту, який розглядає побудову алгоритмів, що можуть навчатися на наявних даних.

**KNN** (K-Nearest Neighbors algorithm) – один з алгоритмів для класифікації об’єктів.

**PNG** (Portable Network Graphics) – один з найбільш популярних растрових графічних форматів, що застосовується для зберігання зображень.

**JPEG** – Joint Photographic Experts Group – один з найбільш популярних растрових графічних форматів, що застосовується для зберігання зображень.

**OpenCV** – бібліотека алгоритмів комп’ютерного зору, обробки зображень та чисельних алгоритмів загального призначення з відкритим кодом.

**Tkinter** – крос-платформна графічна бібліотека на основі засобів Tk, що входить у стандартну бібліотеку Python.

**RFID** (Radio frequency identification) — радіочастотна ідентифікація.

**НМ** – Нейронна Мережа.

**PDF** (Portable Document Format) — формат файлу, створений і підтримуваний компанією Adobe Systems, для представлення двовимірних документів.

**PFD** (Professional Format Disc, професійний диск) — формат оптичного носія.

**TIFF** (англ. Tag Image File Format або Tagged Image File Format) — графічний формат збереження даних.

					ІАЛЦ.045490.004 ПЗ	Анк
Изм.	Лис	№ докум.	Підпис	Дата		2

**DjVu** (від фр. déjà vu — дежавю — колись вже бачене) - компактний формат для зберігання відсканованих документів, книг.

**MP3** — формат файлу для зберігання аудіо-інформації.

**EMV** (Europay, Mastercard, Visa) - міжнародний стандарт, для проведення операцій з банківськими платіжними картками.

					ІАЛЦ.045490.004 ПЗ	Анк
						3
Изм.	Лис	№ докум.	Підпис	Дата		

## ВСТУП

Сьогодні наше життя складно собі уявити без Інтернету. Він є універсальним середовищем для спілкування, розваг та навчання. Із розвитком інформаційних технологій було спрощено багато громіздких процедур. Популярність глобальної мережі Інтернет призвела до виникнення могутнього імпульсу розвитку нових підходів і рішень в найрізноманітніших областях світової економіки. Новим течіям піддалися навіть такі консервативні системи, як системи електронних платежів в банках. Це виразилося в появі і розвитку нових систем платежів – систем електронних платежів через Інтернет.

Платежі через Інтернет – одна з найбільш передових технологій, яка допомагає швидко, ефективно і безпечно вирішити будь-які питання, пов'язані з оплатою рахунків або послуг. Здійснюючи безготівкові платежі, людина може отримати не тільки корисні переваги, але й істотно зекономити власний час і кошти.

З ростом популярності інтернет-платежів, задача спрощення вводу реквізитів банківської карти при проведенні оплати через Інтернет є дуже потрібною.

Розпізнавання даних з банківської карти одночасно є високоактуальним і досить цікавим завданням з точки зору алгоритмів. Добре реалізована програма розпізнавання реквізитів пластикових карток може позбавити людину від потреби вводити більшу частину даних вручну при здійсненні інтернет-платежів і платежів в мобільних додатках.

Процес розпізнавання реквізитів із зображення банківської картки вимагає максимальної точності при проведенні платежів,

					ІАЛЦ.045490.004 ПЗ	Анк
						4
Изм.	Лис	№ докум.	Підпис	Дата		

адже це робота з грошима, і мінімальна помилка може призвести до неприємних наслідків.

Навіть після десятиліть розвитку, розпізнавання даних із зображення не гарантує стовідсоткову точність через ряд непередбачуваних факторів, таких як неякісне зображення, зображення під кутом, незрозумілий фон, з якого важко дістати інформацію, або взагалі відсутність інформації як такої.

Метою даної роботи є розробка алгоритму і програми, що дозволить максимально точно розпізнавати реквізити банківських карток за заданими зображеннями.

					ІАЛЦ.045490.004 ПЗ	Анк
						5
Изм.	Лис	№ докум.	Підпис	Дата		



# 1. АНАЛІЗ ІСНУЮЧИХ РІШЕНЬ ТА ОБГРУНТУВАННЯ ТЕМИ ДИПЛОМНОГО ПРОЕКТУ

## 1.1 Розпізнавання образів

В наш час завдяки використанню сучасних технологій створено нові методи обробки зображень і розпізнавання образів. Але процес створення кожного нового додатку є творчою задачею і потребує додаткових досліджень в зв'язку із специфічними вимогами по розмірності, швидкодії, надійності розпізнавання і об'єму пам'яті, якими характеризується кожна конкретна задача.

Теорія розпізнавання образів існує як розділ інформатики та суміжних дисциплін, що розвиває методи класифікації та ідентифікації об'єктів різної природи: сигналів, ситуацій, предметів, що характеризуються вичерпною кількістю деяких ознак.

Проблема розпізнавання об'єктів також виділена в розділ міждисциплінарних досліджень - в тому числі включаючи роботу зі створення штучного інтелекту, а також часто використовується при вирішенні практичних завдань у галузі комп'ютерного зору.

При постановці класичної задачі розпізнавання об'єктів зазвичай прийнято застосовувати математичну мову, ґрунтуючись на логічних міркуваннях і математичних принципах.

Можна виділити два основних напрямки розпізнавання образів:

- вивчення здібностей до розпізнавання, якими володіють живі істоти, пояснення та моделювання їх;
- розвиток теорії та методів побудови пристроїв, призначених для розв'язку окремих задач в прикладних цілях.

					ІАЛЦ.045490.004 ПЗ	Анк
						6
Изм.	Лис	№ докум.	Підпис	Дата		

Проблема розпізнавання об'єктів також виділена в розділ міждисциплінарних досліджень - в тому числі включаючи роботу зі створення штучного інтелекту, а також часто використовується при вирішенні практичних завдань у галузі комп'ютерного зору.

Розпізнавання образів є однією з найважливіших проблем теорії інтелектуальних систем. Також задача розпізнавання образів має величезне практичне значення. Замість терміну "розпізнавання" інколи використовується інший термін — "класифікація". Часто ці терміни розглядаються як синоніми, але вони не є повністю взаємозамінюваними. Кожний з них має свої сфери застосування, і їх використання залежить від специфіки конкретної задачі.

Класом називають множину об'єктів, які мають спільні властивості. Об'єкти одного класу об'єднуються за наявності «схожості», відповідно до певного опису властивостей цих об'єктів. Для розв'язання задачі розпізнавання визначається довільна кількість класів, але більше одного. Позначається кількість класів числом  $S$ . Кожен з них повинен мати ідентифікуючу мітку свого класу.

Процес призначення міток класу його об'єктам називається класифікацією. Іструментом для присвоєння міток класам є класифікатор. Він отримує перелік ознак об'єкта в якості вхідних даних. Спосіб класифікації, який потребує в описі об'єкта характеристики з набором числових або нечислових ознак, є одним з найпоширеніших. Однак є такі типи даних, для яких відкриті ознаки не дають високої точності класифікації. Прикладом таких даних є звуковий сигнал або колір точок зображень.

Простір ознак - це  $N$ -вимірний простір.  $N$  — це фіксоване число ознак, що були визначені для будь-яких об'єктів.  $N$ -вимірний простір визначений для даної задачі розпізнавання. Вектор з простору ознак  $x$ ,

					ІАЛЦ.045490.004 ПЗ	Анк
						7
Изм.	Лис	№ докум.	Підпис	Дата		

відповідний об'єкту задачі розпізнавання. Це N-вимірний вектор з компонентами  $(x_1, x_2, \dots, x_N)$ , які утворюють значення ознак для даного об'єкта .

Отже, розпізнаванням образів - це групування вихідних даних за їх характеристиками до певного класу. Здійснюється воно за допомогою виділення істотних ознак або властивостей із загальної маси несуттєвих деталей. Якщо вихідним матеріалом служить отримане із камери зображення, то завданням є одержання векторів, що складаються з ознак для кожного класу на зображенні. Даний процес можна розглядати як кодування.

## 1.2 Розпізнавання тексту

Завдання розпізнавання текстової інформації на зображеннях є дуже важливим. Ціль такого проекту – це автоматизація обробки документів, впровадження непаперових технологій. Така задача є однією з найважчих, адже відноситься до повністю автоматичного аналізу зображень. Читаючи рукописний текст, людина помиляється в середньому на 4%.

Розпізнавання тексту або оптичне розпізнавання символів (OCR) – це процес перекладу графічного зображення рукописного чи друкованого тексту в текстові дані – кодовий набір, для представлення символів в комп'ютері[1].

Основна задача оптичного розпізнавання:

- переклад в електронний формат наукових публікацій, документів, історичних видань та соціальної інформації ;
- накопичення і зберігання електронних документів;
- класифікація документів.

Розпізнавання дуже спрощує процес пошуку даних, редагування, різних типів документів, таких як: PDF-файли або знімки, зроблені цифровою камерою, відскановані паперові документи. Наприклад, у

					ІАЛЦ.045490.004 ПЗ	Анк
						8
Изм.	Лис	№ докум.	Підпис	Дата		

бібліотеках тисячі книг та документів регулярно відскановують для резервного копіювання та архіву. Незручністю обробки та зберігання відсканованих документів в таких великих обсягах є нездатність пошуку конкретної фрази або імені всередині файлу. Також текст не можливо скопіювати, підкреслити або змінити, оскільки документ містить великі зображення, а не окремі текстові символи.

Вирішенню багатьох наукових та прикладних задач при ідентифікації об'єктів різної природи допомагають сучасні методи розпізнавання символів. Наприклад, вони використовуються для розпізнавання тексту (вирішення типових задач), або орієнтованих на розпізнавання символічної інформації, нанесеної на поверхню різних об'єктів (спеціалізованих задач).

В загальному випадку, для розпізнавання тексту використовують такі методи:

- розпізнавання з використанням підготовленого шаблону для порівняння;
- структурні методи розпізнавання;
- ознакові методи розпізнавання;
- розпізнавання за допомогою нейронних мереж.

Шаблонні методи перетворюють зображення окремого символу в растрове. Вони порівнюють його зі всіма шаблонами, що є в базі, і серед всіх вибирають шаблон, що містить найменшу кількість крапок, відмінних від вхідного зображення. Такі методи мають високу швидкість обробки вхідних даних та стійкі до різних дефектів зображення. Проте шаблонні методи швидко та якісно розпізнають тільки ті символи, шаблони яких їм „відомі”. Наприклад, якщо шрифт, який потрібно розпізнати, хоч трохи

					ІАЛЦ.045490.004 ПЗ	Анк
						9
Изм.	Лис	№ докум.	Підпис	Дата		

відрізняється від еталонного, такі методи можуть робити помилки навіть при обробці дуже якісних зображень.

Структурні методи розпізнавання зберігають інформацію про його топологію, а не про поточкове написання символу. Перевага структурного методу – стійкість до зсуву і повороту символу на невеликий кут, до різних стильових варіацій шрифтів. Однак цей метод містить недоліки. Основною його проблемою є ідентифікація знаків, які містять певні дефекти: розрив ліній або з'єднання сусідніх ліній. Даний метод не використовують для розпізнавання символів при повороті на кут, який більший десяти градусів. При застосуванні структурного методу важливими вважаються такі ознаки: розмір букв, що розпізнаються, і навіть шрифт, яким вони надруковані.

Розпізнавання за допомогою ознакового методу полягає в тому, що кожному зображенню у відповідність ставиться N-мірний вектор ознак. Такі методи зосереджені на порівнянні вектора ознак з набором еталонних векторів тієї ж розмірності. Ознаковий метод має такі переваги: висока швидкість розпізнавання, нескладність реалізації, хороша узагальнююча здатність. Цей метод має і свої недоліки. На етапі, коли відбувається виділення ознак про символ, проходить незворотня втрата частини інформації. Під час використання цього методу відбувається незалежне виділення ознак, тому інформація про взаємне розташування елементів символів втрачається. Крім того, є висока чутливість до дефектів зображення.

### **1.3 Розпізнавання реквізитів банківської картки**

Завдання розпізнавання даних із банківської карти за фотографією пов'язане з двома дисциплінами: цифровою обробкою та штучним інтелектом. Є ряд суттєвих проблем, пов'язаних із мінливістю оточуючих параметрів: освітленості, кольору, масштабу, ракурсу спостереження і т. д.

					ІАЛЦ.045490.004 ПЗ	Анк
						10
Изм.	Лис	№ докум.	Підпис	Дата		

Яскравість та колір пікселів на зображенні також залежать від великої кількості важко прогнозованих факторів:

- джерела світла (кількості та його розташування);
- кольору та інтенсивності випромінювання;
- тіні та віддзеркалення світла від оточуючих об'єктів.

З точки зору розпізнавання банківська картка – це складний стандартний документ, виконаний на типовому бланку і містить певний набір інформаційних полів (обов'язкових і додаткових). Частина полів знаходиться на лицьовій стороні, друга частина – на зворотній.

Виготовлена банківська платіжна картка із спеціального стійкого до механічних пошкоджень пластику відповідно до стандарту ISO/IEC 7810. Це міжнародний стандарт, який має такі фізичні характеристики: стійкість до згинання, хімікатів, температури, вогню і вологості. Для платіжних карток використовується формат ID-1. Він визначає розмір  $85,60 \times 53,98$  мм і заокруглені кути. Картка має товщину 0,76 мм. Ідентифікування забезпечується нанесенням на картку її номера, терміну дії, прізвища користувача, ім'я користувача і зразка підпису або інших ідентифікаційних даних.

В загальному випадку платіжна картка складається із наступних елементів[8]:

- банківський логотип;
- EMV-чіп (EMV -схема, інтегральна схема);
- голограма;
- номер картки (банківський номер, платіжний номер);
- логотип платіжної системи;
- термін дії;

					ІАЛЦ.045490.004 ПЗ	Анк
						11
Изм.	Лис	№ докум.	Підпис	Дата		

- ім'я отримувача картки (ім'я користувача картки (юридично, власником карти являється банк)).

Для здійснення інтернет-платежів в більшості випадків потрібно розпізнати на зображенні номер картки, ім'я власника і термін дії банківської картки.

В загальному випадку алгоритм вирішення задачі розпізнавання трьох цільових полів за зображенням банківської картки складається з наступних кроків[10]:

1. Встановлення факту наявності банківської картки на зображенні.
2. Попередня фільтрація зображення (з метою приглушення фону банківської картки, який буває досить насиченим).
3. Пошук зон (рядків) цільових інформаційних полів на зображенні.
4. Сегментація віднайдених рядків на “коробки символів”.
5. Розпізнавання віднайдених “коробок символів”.
6. Застосування різних методів перевірки отриманих результатів.

В залежності від конкретних умов структура та реалізація окремих кроків алгоритму можуть відрізнятися. Розпізнавання інформаційних полів на банківській картці може мати різну складність і в залежності від цього склад кроків дуже змінюється. Найлегше розпізнати номер картки, яка відповідає наступним вимогам :

- номер банківської картки складається тільки із цифр;
- формат номера строго визначений для кожного типу платіжної картки;
- геометричне положення номера на картці більш стабільне незалежно від виробника;
- існує алгоритм Луна, який допомагає визначити правильність розпізнавання номера банківської картки.

Набагато складніше розпізнати такі реквізити платіжної картки, як термін дії та ім'я користувача. Найважчим реквізитом, який потрібно розпізнати, є термін дії картки. На відміну від номера платіжної картки, розміщення цього поля не є стандартним. Використання методу “грубої сили” (тобто пошуку по всій площині картки) не дає перспективних результатів. Імовірність помилкового виявлення на довільному фрагменті тексту чи навіть строкатій ділянці фону є дуже великою.

#### 1.4 Аналіз існуючих рішень

Перетворенням графічного зображення в текст займаються спеціальні програми розпізнавання тексту (Optical Character Recognition - OCR)[1].

Сучасна OCR повинна вміти багато: розпізнавати тексти, набрані не тільки звичайними шрифтами, але і найекзотичнішими. Вона має коректно працювати із текстами, які мають слова на декількох мовах, із рукописними текстами, коректно розпізнавати таблиці. Не менш важливо забезпечити можливість збереження результатів у файл популярного текстового формату.

Найбільш затребувані системи оптичного розпізнавання документів, наприклад, ABBYY FineReader і CuneiForm від Cognitive, використовують як растровий, так і структурний методи розпізнавання. Крім того, ці системи “самонавчальні” (для кожного конкретного документа вони створюють відповідний набір шаблонів символів), тому швидкість і якість розпізнавання багатосторінкового документа поступово зростає.

Компанія Soft Xpansion Ukraine, що є партнером компанії «ABBYY», пропонує розпізнавання документів за допомогою технологій ABBYY. Soft Xpansion Ukraine - це одна з провідних компаній в Україні з розробки та впровадження систем електронного документообігу, систем

					ІАЛЦ.045490.004 ПЗ	Анк
						13
Изм.	Лис	№ докум.	Підпис	Дата		



електронних архівів, порталних рішень, що дозволяють автоматизувати бізнес-процеси зберігання, поширення та обміну інформацією і документами.

ABBYY FineReader - це універсальна програма для оптичного розпізнавання тексту, що дозволяє підвищити продуктивність при роботі з документами. Вона надає потужні, але водночас прості у використанні інструменти для отримання доступу до інформації, що міститься у паперових документах і PDF-файлах. Програма спрощує роботу з паперовими документами і PDF-файлами завдяки поєднанню технології оптичного розпізнавання тексту з базовим функціоналом для роботи з PDF-файлами. FineReader проводить розпізнавання тексту з більш 180 мов, для 38 з них передбачена вбудована перевірка орфографії. Починаючи з версії Professional, розпізнаються іврит, японська, тайська, китайська мови. FineReader відкриває файли графічних форматів (TIFF, JPG, PFD, PNG та ін.) В тому числі DjVu - компактний формат для зберігання відсканованих документів, книг.

Для налаштування потокового сканування та розпізнавання створені серверні версії – ABBYY Recognition Server і ABBYY FlaxyCapture. Вони дозволяють працювати з великою кількістю документів різного виду. Наприклад, платіжні документи, анкети, договори.

З використанням даної технології виникла можливість не тільки розпізнавати та зберігати відскановані документи в редаговані формати. Важливою є можливість витягати з різних документів (договорів, рахунків, анкет, тощо) потрібну інформацію і зберігати її у базі даних. Ці дані можна імпортувати в інші системи, наприклад, в тексти системи електронного документообігу. Це стосується не тільки надрукованого тексту, але і рукописних чисел і букв, а також різних типів позначок: штрих-кодів, будь-яких поміток в клієнтських анкетах. Дана технологія

					ІАЛЦ.045490.004 ПЗ	Анк
						14
Изм.	Лис	№ докум.	Підпис	Дата		

спростила процес ведення документації, збільшила швидкість введення даних. Підвищилась ефективність пошуку документів в електронному архіві. Пошук можна робити за атрибутами документа або в тексті. При цьому гарантується достовірність даних внаслідок високої якості розпізнавання.

Головним конкурентом технологій ABBYY, який займається оцифровкою тексту, є додаток Readiris. Це функціональний інструмент для розпізнавання тексту як зі сканера, так і зі збережених файлів різних форматів (PDF, PNG, JPG і ін.). Основне завдання програми – це розпізнавання тексту, розміщеного у файлах графічного формату: на картинках, в PDF. Вона може розпізнавати текст, що міститься навіть в нестандартних форматах: файлах MP3 або FB2. Крім того, Readiris має унікальну здатність розпізнавати рукописний текст. Хоча за кількістю функцій дана програма дещо поступається ABBYY FineReader, але набагато перевершує більшість інших програм. Головною перевагою Readiris є можливість інтеграції з великим переліком хмарних сервісів для зберігання файлів. Недоліком у Readiris є висока вартість платної версії.

Програма OmniPage відрізняється від своїх аналогів високою швидкістю і точністю розпізнавання. Розпізнаються більше 120 мов з різними алфавітами: латинським, грецьким, кирилицею, а також китайською, японською і корейською мовами. Як і FineReader, OmniPage впевнено розпізнає документи, отримані за допомогою цифрових камер за допомогою технології корекції зображення "3D Correction".

Головне завдання програми VueScan - це сканування документів з паперових носіїв, а не процес розпізнавання тексту. Переваги програми в тому, що вона працює з дуже великим переліком сканерів. Для подальшої взаємодії додатка з пристроєм навіть не потрібна установка драйверів. Ця

					ІАЛЦ.045490.004 ПЗ	Анк
Изм.	Лис	№ докум.	Підпис	Дата		15

програма може працювати з додатковими можливостями сканерів, для яких навіть рідні додатки не можуть забезпечити такі результати.

VueScan містить інструмент розпізнавання сканованого тексту, але він не користується популярністю. Функція оцифрування тексту є незручною і досить слабка. Тому розпізнавання в додатку використовується для вирішення нескладних завдань. Отже, VueScan - відмінний додаток для сканування.

Справляється із завданням оптичного розпізнавання тексту у файлах PDF програма Adobe Export PDF. Вона виконує оптичне розпізнавання тексту у файлах PDF, які створені зі сканованих документів. Тут можуть мати місце зображення, векторна графіка, прихований текст. Програма навіть виконає завдання оптичного розпізнавання тексту, який не вдається прочитати через те, що його неправильно закодовано у вихідній програмі.

Розглянемо роботу та технологію компанії Jumio, що здійснює онлайн-перевірку мобільних платежів, надає послуги зі сканування та перевірки ідентифікаційних даних для проведення онлайн транзакцій. Вона використовує технології комп'ютерного зору, машинного навчання для перевірки документів як, наприклад: паспорт, посвідчення водія, тощо, виданих більш ніж 200 країнами. Технології Jumio, за словами самої компанії, вважаються найбільшими і найточнішими рішенням на ринку і використовується провідними компаніями у сфері фінансових послуг, що розділяють сектори економіки, роздрібної торгівлі, подорожі, блокчейн та онлайн-ігор.

Компанія також має реалізований сервіс, що забезпечує отримання реквізитів банківських карток. Він працює наступним чином: клієнт на сайті хоче ввести реквізити картки за допомогою її фотографії. Для отримання реквізитів банківської карти потрібно перенаправити користувача на веб-сторінку, яку пропонує нам сервіс. Там він слідує

					ІАЛЦ.045490.004 ПЗ	Анк
						16
Изм.	Лис	№ докум.	Підпис	Дата		

інструкціям, що для нього відображаються. Після того, як користувач здійснив всі потрібні кроки інструкції, його повертають на початковий сайт. Одночасно ми отримуємо повідомлення із реквізитами платіжної карти на url-адресу зворотнього виклику, яку ми передали компанії раніше. Недоліки роботи сервісу будуть розглянуті в наступному розділі.

Додаток CuneiForm - відмінне програмне рішення для розпізнавання тексту з фото, файлів зображень, сканера. В ньому застосовується особлива технологія оцифровки. Вона поєднує шрифтове і незалежне від шрифту розпізнавання. Завдяки такому додатку можна дуже точно розпізнавати текст, враховуючи навіть елементи форматування. При цьому швидкість роботи залишається високою. Існують багато платних програм для розпізнавання тексту. Перевагою даного додатку є той факт, що цей додаток абсолютно безкоштовний.

Проте додаток CuneiForm містить ряд недоліків. Він несумісний з деякими моделями сканерів та не працює з одним з найуживаніших форматів - PDF. Ще однією незручністю є той факт, що на даний момент він офіційно не підтримується розробниками.

Оцифровкою отриманого зі сканера тексту в формат PDF займається додаток WinScan2PDF. На відміну від CuneiForm, це є його єдиною функцією. Найбільша перевага даної програми – це простота використання. Додаток буде найбільш затребуваний тими людьми, які дуже часто сканують паперові документи і розпізнають текст в форматі PDF.

Головний недолік WinScan2PDF - це дуже обмежені функції. Даний додаток нічим більше не займається, крім зазначеної вище процедури. Він не має змоги проводити оцифровку файлів зображень, що вже збережені на комп'ютері, не може зберігати результати розпізнавання в інший формат, окрім PDF.

					ІАЛЦ.045490.004 ПЗ	Анк
						17
Изм.	Лис	№ докум.	Підпис	Дата		

Найбільш універсальним офісним додатком для розпізнавання тексту та сканування документів є RiDoc . Його функції не такі обмежені, хоча набагато менші ніж в ABBYY FineReader або Readiris. Але по співвідношенню «ціна - якість» додаток має переваги, адже вартість даного продукту набагато менша. Обмеження по функціоналу в даній програмі не є істотними. RiDoc добре справляється із завданнями сканування та розпізнавання. Перевагою його є можливість зменшення зображень без втрати якості.

Єдиний істотний недолік програми - не зовсім коректна робота з розпізнавання дрібного тексту.

Будь-який користувач зможе відшукати серед перерахованих програм той додаток, який йому найбільше підходить. Його вибір залежатиме як від поставлених завдань, які користувачеві доводиться найчастіше вирішувати, так і від його фінансового забезпечення.

### **1.5 Обґрунтування теми дипломного проекту**

Друкування тексту для приведення його в електронний вигляд давно вже відійшло у минуле. Адже наявність досить потужних системи для розпізнавання цифрової інформації значно спрощує життя людини. Робота з такими системою вимагає мінімального втручання користувача. Тож програми для оцифровки тексту затребувані як в офісі, так і вдома.

На сьогоднішній день існує ряд методів, що вирішують проблему розпізнавання друкованого тексту, частину з них ми розглянули вище. Однак розпізнавання банківської картки - це вузька спеціалізована область, що вимагає максимальної точності. Незважаючи на це, можна підібрати метод, який буде максимально точно і швидко розпізнавати реквізити банківської картки. Ми розглянули вище програмні рішення, що займаються розпізнаванням символів, але ці програми використовують

					ІАЛЦ.045490.004 ПЗ	Анк
						18
Изм.	Лис	№ докум.	Підпис	Дата		

методи розпізнавання, які показують чудові результати на однорідному, в більшості випадків, білому фоні. Зображення площини банківської платіжної картки може бути як завгодно різноманітне.

На даний момент існує велика кількість програмних рішень для розпізнавання символів із зображення. Але серед них не є багато таких, які займаються розпізнаванням документів, в нашому випадку банківської карти.

Під час аналізу існуючих рішень було розглянуто технології компанії Jumio, яка вражає своїми потужними можливостями, але все ж таки інтеграція такого сервісу має ряд значних недоліків:

- сервіс не працює із ембосованими банківськими платіжними картками;
- технологія, що використовується, не є безпечною, адже дані банківської картки передаються напряму без відповідного шифрування;
- використання даної послуги потребує чималих фінансових витрат.

Програмні методи не завжди є достатньо ефективними – більшість систем орієнтовані на використання одного з методів розпізнавання символів паперових носіїв, що не завжди підходить для різнорідних вхідних даних, які можуть мати низьку швидкість розпізнавання або малу якість розпізнання символів.

Отже, існує проблема недосконалості сучасних методів розпізнавання символів, що не забезпечують високу якість та швидкість розпізнавання. Тому необхідно розробити метод, який буде мати більшу продуктивність розпізнання символів та швидкість перетворення паперових документів в електронний варіант.

					ІАЛЦ.045490.004 ПЗ	Анк
						19
Изм.	Лис	№ докум.	Підпис	Дата		

Таким чином, дане дослідження полягає у аналізі існуючих методів розпізнання символів та розробці власного методу, який буде забезпечувати достатню якість та швидкість розпізнавання реквізитів банківської картки із найрізноманітнішим фоном із зображення. Необхідно розробити програму, яка дозволить за допомогою зручного інтерфейсу завантажувати зображення або моментально робити фотографію банківської картки для подальшої обробки, обрати зручний спосіб представлення та збереження результатів обробки.

					ІАЛЦ.045490.004 ПЗ	Анк
						20
Изм.	Лис	№ докум.	Підпис	Дата		

## **2 ХАРАКТЕРИСТИКА ІНСТРУМЕНТАРІЮ ТА ОПИС ПРОЦЕСУ РОЗРОБКИ ПРОГРАМИ**

### **2.1 Методи та алгоритми попередньої обробки зображень**

Застосування цифрової техніки підсилює розвиток нових цифрових методів обробки сигналів. Пришвидшує даний процес впровадження нових сучасних комп'ютерних та телекомунікаційних технологій. На сьогоднішній день великого значення набувають методи цифрової обробки зображень.

Область обчислювальної техніки, що охоплює технічні і програмні засоби, називається цифровою обробкою зображень.

Ідентифікацію досліджуваних об'єктів здійснюють у рамках аналізу зображень: розпізнавання образів та аналіз сцен. Також оцінюють взаємозв'язок окремих фрагментів зображення, визначають характеристики зображених об'єктів.

В даному процесі можна виділити дві групи методів, що використовуються для обробки зображень:

- поелементна;
- кореляційна.

Результат обробки будь-якої точки кадру зображення, за умови використання методів першої групи, залежить загалом від значення відліку потрібного параметра первинного зображення у цій же точці. Доступність і простота реалізації є основною перевагою таких процедур. Використання методів цієї групи призводить до значного суб'єктивного покращення візуальної якості. Заключним етапом процесу комплексної обробки зображень є поелементна обробка зображень..

Можна розділити процедури поелементної обробки на:

- процедури пов'язані зі зміною контрасту зображення;

					ІАЛЦ.045490.004 ПЗ	Анк
						21
Изм.	Лис	№ докум.	Підпис	Дата		



– процедури бінаризації зображень.

Одним з найпоширеніших методів виділення об'єктів на зображеннях є сегментація. Вона доволі часто застосовується для виділення областей, однорідних за деякою складною властивістю, що називаються кластерами.

Найпопулярнішими способами сегментації зображень є три основні: порогова сегментація зображень, сегментація шляхом нарощування та шляхом виділення кордонів.

Перетворення з пороговою характеристикою називається бінаризацією. Результатом бінаризації є те, що напівтоноване зображення, яке містить всі рівні яскравості, перетворюється у таке, що містить тільки два рівні яскравості  $y = 0$  або  $y = u_{\max}$ .

Коли основною метою обробки зображення є виділення контурів об'єктів, застосовують операцію бінаризації. Визначення порогового значення є основною проблемою цього методу, так як порівняння яскравості вхідного зображення із пороговим дозволяє визначити значення кінцевого зображення в кожній його точці.

Значення кольору є порогом для зображень в градаціях сірого та RGB зображень. В більшості випадків вибір порогу при бінаризації здійснюється за допомогою адаптивного алгоритму. Вибір математичного очікування, моди або піків гістограми є одними з найдоцільніших алгоритмів. Бінаризація ефективна для сегментації кольорів при роботі з гістограмою.

При компресії зображень застосовують одномірне перетворення Фур'є. Але для обробки зображень перетворення Фур'є майже не використовується в чистому вигляді. Причиною цього є те, що одновимірне перетворення зазвичай не вистачає для аналізу зображень.

					ІАЛЦ.045490.004 ПЗ	Анк
Изм.	Лис	№ докум.	Підпис	Дата		22

Тоді застосовують більш ресурсне двовимірне перетворення. При аналізі спектра доцільно використовувати саме цей метод.

Етап попередньої обробки зображень є одним з найважливіших під час розпізнавання символів із зображення. Адже результатом виконання даного етапу є приглушення фону зображення, виділення контурів об'єктів, які будуть розглядатися на наступних етапах.

## **2.2 Штучний інтелект**

Одним з найперспективніших напрямків комп'ютерних наук, який вивчає методи розв'язання задач, для яких не існує првильних способів вирішення є штучний інтелект. Системи, які використовують штучний інтелект можуть самонавчатися та оперувати даними. Застосування таких систем в різних сферах є необмеженим. Наприклад, створення роботів, що можуть самостійно приймати рішення, машини з автопілотом чи перекладачі онлайн в реальному часі[9].

Штучний Інтелект використовується в багатьох сферах: робототехніка, комп'ютерні ігри, веб-аналіз, прийняття рішень, обробка зображень, прогнозування, медицина. Штучний інтелект, який застосовується для розпізнаванні образів дозволяє створювати майже працюючі системи розпізнавання графічних об'єктів, використовуючи аналогічні ознаки. Ознаками можуть бути будь-які характеристики об'єктів, які можна розпізнати. Вони повинні бути інваріантні до розташування, форми та розміру об'єктів. Одна з основних задачею СШІ є сегментація та визначення об'єктів на зображеннях. Звуковий набір та рукописний текст в мобільних телефонах, визначення локації будинку та промальовування його внутрішньої структури, також є прикладами штучного інтелекту.

Ще в середині ХХ століття відбулися перші дослідження з вивчення штучного інтелекту. Поняття штучного інтелекту як науки першим

					ІАЛЦ.045490.004 ПЗ	Анк
						23
Изм.	Лис	№ докум.	Підпис	Дата		

сформулював професор Джон Маккарті у Дартмутського коледжу в 1956 році, а вершиною цих наукових пошуків став один з відомих тестів - тест Тьюринга.

Хоча існують суперечки про те, чи можна відносити ІІ до інтелекту повною мірою. Вони продовжуються й до цього часу. Та це не заважає тому, що розумні системи, які можуть навчатися, дуже полегшують життя людини.

Основні переваги використання штучного інтелекту[12]:

- точність в обробці даних;
- здатність аналізувати велику кількість інформації з великою швидкістю;
- ІІ не потрібен сон і перерва на обід, він не допускає помилок через перевтому;
- використовувати штучний інтелект можна там, де людині небезпечно перебувати.

Використання машин, які можуть навчатися, значно економить час та фінансові витрати і сприяє більшій продуктивності праці.

### 2.3 Машинне навчання

Машинне навчання – це потужний підрозділ штучного інтелекту, який розглядає побудову алгоритмів, що можуть на наявних даних навчатися. Розглянемо задачу машинного навчання. Існує певний набір об'єктів – прикладів і певний набір відповідей. Між ними існує певна прихована залежність. Завдання машинного навчання заключається в тому, щоб для прогнозування відповідей на основі нових даних знайти приховану залежність. Для даного навчання використовують різні технології та алгоритми. Серед них є такі: байєсовські класифікатори, дискримінантний аналіз та багато інших математичних методів.

Машинне навчання поділяється на 3 типи:

					ІАЛЦ.045490.004 ПЗ	Анк
						24
Изм.	Лис	№ докум.	Підпис	Дата		

1) навчання з вчителем – існує набір прикладів і є правильна відповідь до кожного прикладу. Потрібно по прикладах навчитися надавати правильну задану вчителем відповідь.

2) навчання без вчителя – існує великий набір даних, які містять приховані закономірності. Потрібно ці закономірності знайти. Для цього можна розбити дані на кластери чи певні групи.

3) навчання з підкріпленням – полягає у взаємодії програми з динамічним середовищем. У ньому програма повинна виконувати завдання без учителя, який вказував би їй, чи близько вона наблизилась до мети. Прикладом такого навчання є навчання гри через гру.

Машинне навчання класифікуються ще за видом розв’язуваної проблеми :

- задача класифікації – віднесення об’єкту до однієї з категорій на основі її характеристик. Це задача навчання з вчителем. Для неї потрібні набори даних: навчальна та тестова вибірки.

- задача кластеризації – розбиття на основі характеристик множини об’єктів на групи так, щоб в одній групі були схожі між собою об’єкти;

- задача регресії – прогнозування на основі інших характеристик об’єкта його кількісної характеристики;

- задача виявлення аномалій – пошук у вибірці об’єктів, «дуже не схожих» на всі інші або на якусь групу об’єктів .

Найбільш використовувані методи машинного навчання для вирішення задачі класифікації – це дерева рішень та метод найближчих сусідів.

Дерево рішень – це елементарний класифікатор, який є об’єднанням логічних правил в структуру даних «дерево». Даний метод є достатньо наочним. Тобто, дерева рішень можуть легко візуалізуватись та зображувати конкретне рішення. При невеликій кількості параметрів є

					ІАЛЦ.045490.004 ПЗ	Анк
Изм.	Лис	№ докум.	Підпис	Дата		25

досить швидким процес навчання та прогнозування. Перевагою методу є підтримка категоріальних та числових показників. Недоліком є чутливість до навчальних даних. Якщо навчальна вибірка трохи зміниться, може повністю змінитись дерево. Для боротьби з перенавчанням потрібно відсікати гілки дерева або встановлювати максимальну глибину дерева чи мінімальне число елементів в листках дерева. Також існує проблема пошуку оптимального дерева рішень

Метод найближчих сусідів ( $k$  Nearest Neighbors, або  $kNN$ ) – це відносно простий метричний алгоритм для автоматичної класифікації об'єктів. Він був запропонований у 1973 р. Передбачається, що існує певний набір об'єктів з уже наявною класифікацією. Кожен об'єкт має кінцеву кількість розмірностей. В основі даного алгоритму є об'єкти, які є близькими за значеннями  $n-1$  властивості (ознаки), тоді вони близькі за значеннями  $n$ -ї ознаки. Відстань між двома класами – це відстань між найближчими їх об'єктами.

Спочатку між об'єктами розраховується матриця відстаней. В даній матриці на кожному кроці відбувається пошук мінімального значення. Воно співпадає з відстанню між двома найбільш близькими кластерами. Знайдені кластери потрібно об'єднати. Створюється новий кластер, який об'єднує знайдені кластери. Поки не будуть об'єднані всі кластери, процедура повторюється.

Кластер, відповідно до методу  $kNN$ , утворюється за правилом: об'єкт, що знаходиться дуже близько хоча б до одного елемента даного кластера, потрапляє до вже існуючого кластера. Отже, приєднання залежить лише від зв'язку між кластером і об'єктом.

					ІАЛЦ.045490.004 ПЗ	Анк
						26
Изм.	Лис	№ докум.	Підпис	Дата		

Математичні властивості є важливою перевагою данного методу: отримані за цим методом результати не залежать від монотонних перетворень матриці схожості.

Недолік методу полягає в тому, що з'являються "ланцюжки". Тобто утворюються великі продовгуваті кластери. Особливу увагу при використанні методу kNN потрібно звернути вибору міри відстані між об'єктами. Адже на підставі міри відстані формується первинна матриця відстаней. Подальший процес класифікації визначає саме вона.

## **2.4 Штучна нейронна мережа**

В кінці XX століття велика увага стала приділятися штучним нейронним мережам. Перша модель нейронної мережі була запропонована в 1957 році. Вона реалізувала алгоритми машинного навчання, схожі на сучасні. Черговий вибух інтересу до нейронної мережі почався в 1986 році, після важливого розвитку методу зворотного поширення помилки, який застосовувався при навчанні нейронної мережі.

Штучна нейронна мережа – це модель біологічної нейронної мережі. Вона складається з шарів штучних нейронів, які різним чином пов'язані між собою. Штучний нейрон складається з двох частин — адаптивного суматора та нелінійного перетворювача. Нейрони функціонально впливають на роботу один одного і організовують загальну активну структуру. Активність нейрона у більшості архітектур штучних нейронних мереж визначається перетворенням зовнішнього сумарного впливу інших нейронів на даний нейрон.

Технології штучних нейронних мереж на етапі свого зародження розвивалися досить відокремлено від класичних методів. Вони залишали значний вплив на теоретичний, термінологічний і методологічний апарати теорії машинного навчання і розпізнавання об'єктів. Ці технології інколи

					ІАЛЦ.045490.004 ПЗ	Анк
Изм.	Лис	№ докум.	Підпис	Дата		27

докорінно змінювали уявлення про предмет і проблематику цих дисциплін. Розвиток базових моделей штучних нейронних мереж дав поштовх до поділу науки про нейронні мережі на види: топології архітектури мереж і методи навчання мереж.

Для більшості архітектур штучних нейронних мереж функції активації нейронів є фіксованими, ваги синапсів виступають в ролі параметрів мережі. Основним завданням нейронної мережі є перетворення вхідного вектора у вихідний вектор. Воно здійснюється вагою і топологією мережі. Зовнішніми входами сукупної мережі є деякі входи нейронів, а виходами сукупної мережі – деякі виходи нейронів.

Можна характеризувати штучний нейрон за його поточним станом.

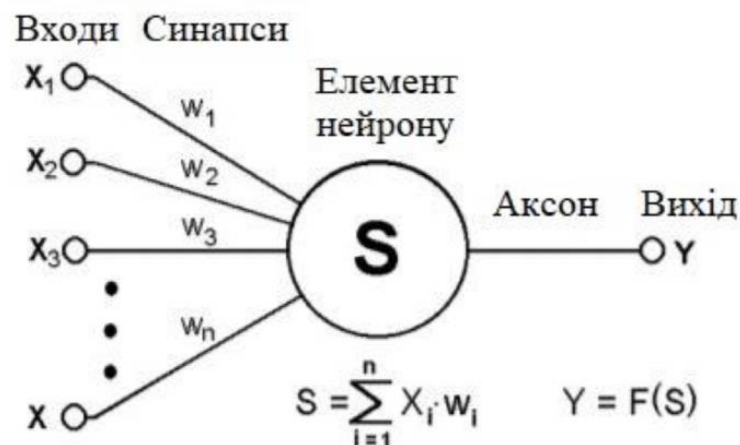


Рисунок 1 – Модель штучного нейрона

Він володіє групою синапсів. Це зв'язки, по яких вихідні сигнали перших нейронів надходять на входи других. Кожен такий зв'язок характеризується своєю вагою. Одиначний вихід нейрона називають аксоном. З нього сигнал (збудження або гальмування) надходить на синапси наступних нейронів. З одиначного виходу нейрона сигнал може надходити на довільне число входів до інших нейронів. Це односпрямовані вхідні зв'язки, які з'єднані з виходами інших нейронів. Загальну модель в штучного нейрона наведено на рисунку 1.

Доречна аналогія з нервовими клітинами головного мозку. Входи  $x_1$ – $x_n$  відповідають дендритам нейрону. Саме до них надходять сигнали через синапси, які розділяються на збуджуючі та гальмуючі, що відповідає добутку вхідних сигналів  $x_1$ – $x_n$  на вагові коефіцієнти  $w_1$ – $w_n$ . При цьому, якщо ваговий коефіцієнт має знак плюс, то даний вхід збуджуючий, якщо ж мінус — то гальмуючий. Далі сигнали потрапляють на суматор та функцію активації, що в природі відповідає складенню сигналів в тілі нейрона та відправлення певного сигналу по аксону.

Від вибору функції активації залежить не тільки швидкість, але й метод навчання нейронної мережі. Останнє пов'язане з тим, що відомі алгоритми навчання накладають певні обмеження на функцію активації. Більшість функцій активації мають стискаючі властивості, тобто вихідні значення нейрона завжди належать певному інтервалу, який залежить від типу функції активації. В теорії побудови нейронних мереж використовують велику кількість функцій активації. Серед них слід виокремити чотири, які використовуються найчастіше.

Перший штучний нейрон мав функцію активації у вигляді порогової функції (інша назва – функція Хевісайда). Вона найкраще відповідала характеристикам біологічного нейрону: значення виходу нейрона рівне нулю до тих пір, доки на виході суматора не буде значення, яке перевищує пороговий рівень. Як тільки це сталося – нейрон переходить в збуджений стан і на виході з'являється одиниця. До переваг функції слід віднести простоту опису та швидкість розрахунків. З недоліків є непридатність для роботи з неперервними сигналами (вихід приймає значення або 0, або 1), а відсутність першої похідної не дозволяє використовувати деякі з методів навчання.

Лінійна функція з насиченням, на відміну від функції Хевісайда, має проміжні значення в діапазоні від 0 до 1, що дає змогу більш широкого

					ІАЛЦ.045490.004 ПЗ	Анк
						29
Изм.	Лис	№ докум.	Підпис	Дата		



застосування в системах класифікації образів, також таку функцію інтерпретують, як апроксимаційну характеристику нелінійного підсилювача. Недоліком залишається відсутність першої похідної.

Найбільш поширеними передавальними функціями є сигмоїдальні функції активації. Недостатня гнучкість класифікаторів на основі порогових передавальних функцій зумовила пошук потрібної для вирішення цієї проблеми. Сигмоїдальні функції активації відносяться до стискаючих нелінійних функцій. Сигмоїд - це монотонно зростаюча всюди диференційована S-образна нелінійна функція з насиченням. Введення нелінійності в роботу нейронної мережі дозволяє будувати ефективні багатошарові нейронні мережі, які здатні до апроксимації складних функцій. На відміну від порогової, сигмоїдальні функції диференційовані на всій числовій осі. Вони забезпечують посилення слабких сигналів і запобігають насиченню сильних сигналів, тому часто використовуються в нейронних мережах. Використання сигмоїдальних функції (1) дозволило перейти від дискретних сигналів до неперервних.

$$f(s) = \frac{1}{1 + e^{-as}} \quad (1)$$

До сигмоїдальних функцій відноситься логістична функція.  $a$  – коефіцієнт крутизни логістичної функції. Вона, в залежності від значення  $a$ , приймає різний кут нахилу, що можна використовувати для регулювання підсилення слабких сигналів та зміни швидкості навчання мережі. До переваг функції відноситься зручне вираження першої похідної через значення функції, що дозволяє ефективно використовувати алгоритм зворотного поширення помилки під час навчання.

Крім логістичної функції до сигмоїдальних відноситься і гіперболічний тангенс, який на відміну від попередньої функції має

					ІАЛЦ.045490.004 ПЗ	Анк
						30
Изм.	Лис	№ докум.	Підпис	Дата		

діапазон вихідних значень, який лежить в інтервалі  $(-1;1)$ , що дозволяє нейронній мережі працювати не лише з додатними, але й з від'ємними величинами .

Функція ReLU (Rectified Linear Unit) є випрямленою лінійною функцією і на даний момент вважається більш простішим і ефективним з точки зору обчислювальної складності варіантом передавальної функції. Похідна цієї функції дорівнює або 0 або 1, від чого її застосування запобігає розростанню і загасанню градієнтів, і призводить до зменшення ваг, що позитивно позначається на обчислювальній здатності нейромережі. Передавальна функція ReLU (2) є одним з останніх успіхів в області методів налаштувань глибоких нейронних мереж.

$$f(x) = \max(0, x) \quad (2)$$

У формулі (2)  $x$  - вхід нейрона.

Сьогодні існує сімейство різних модифікацій ReLU, які вирішують проблеми надійності цієї функції при проходженні через нейрон великих градієнтів: Leaky ReLU, Parametric ReLU, Randomized ReLU. Одношаровою є мережа, що складається з сукупності нейронів, утворення яких формує шар, як зображено на рисунку 2.

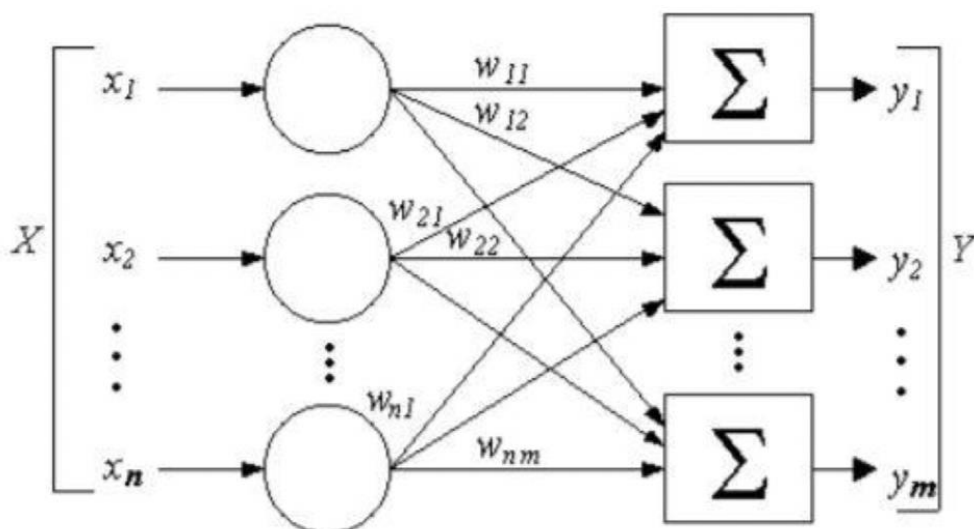


Рисунок 2 - Одношарова штучна нейронна мережа

Ліві вершини схеми зображеного нейрона служать для розподілу сигналів, що подаються на вхід. За цими вершинами не закріплено жодних обчислень, тому вони не вважаються шаром і позначені колами, щоб відрізнити їх від обчислювальних нейронів, що позначаються квадратами. Існує сполучення окремою вагою кожного елементу з множини входів  $X$  з кожним штучним нейроном. За нейроном закріплена робота подачі зваженої суми входів в мережу. З метою спільності штучні та біологічні мережі мають відсутні з'єднання. Існують з'єднання між виходами і входами елементів в самому шарі. Ваги представлені елементами матриці  $W$ . Матриця має  $n$  рядків і  $m$  стовпців, де  $n$  - число входів, а  $m$  - число нейронів. Наприклад,  $w_{12}$  - це вага, що зв'язує перший вхід з другим нейроном. Таким чином, обчислення вихідного вектора  $Y$ , зводиться до матричного множення  $Y = XW$ .

В якості функції активації іноді використовують експоненціальну, модульну, квадратичну, радіально-базисну та інші функції. Вибір функції активації та її параметрів залежить від діапазону вхідних та вихідних значень мережі і методу навчання.

Відмінності обчислювальних процесів в нейронних мережах часто обумовлені способом взаємозв'язків нейронів. За сукупністю критеріїв на сьогоднішній день багат шарові архітектури можна розділити на статичні і динамічні. Кожен з класів архітектур нейронних мереж може включати безліч підкласів, реалізуючи різні підходи.

Мережі прямого поширення, в яких реалізована одно напрямлений зв'язок між шарами, відносяться до статичних архітектур. В них відсутні динамічні елементи і зворотній зв'язок, а вихід навченої нейромережі не залежить від попередніх станів мережі і однозначно визначається входом в мережу.

					ІАЛЦ.045490.004 ПЗ	Анк
Изм.	Лис	№ докум.	Підпис	Дата		32

Приклади статичних штучних нейронних мереж прямого поширення:

- Персептрон
- Нейронна мережа Кохонена
- Когнітрон та неокогнітрон
- Сучасна згорткова

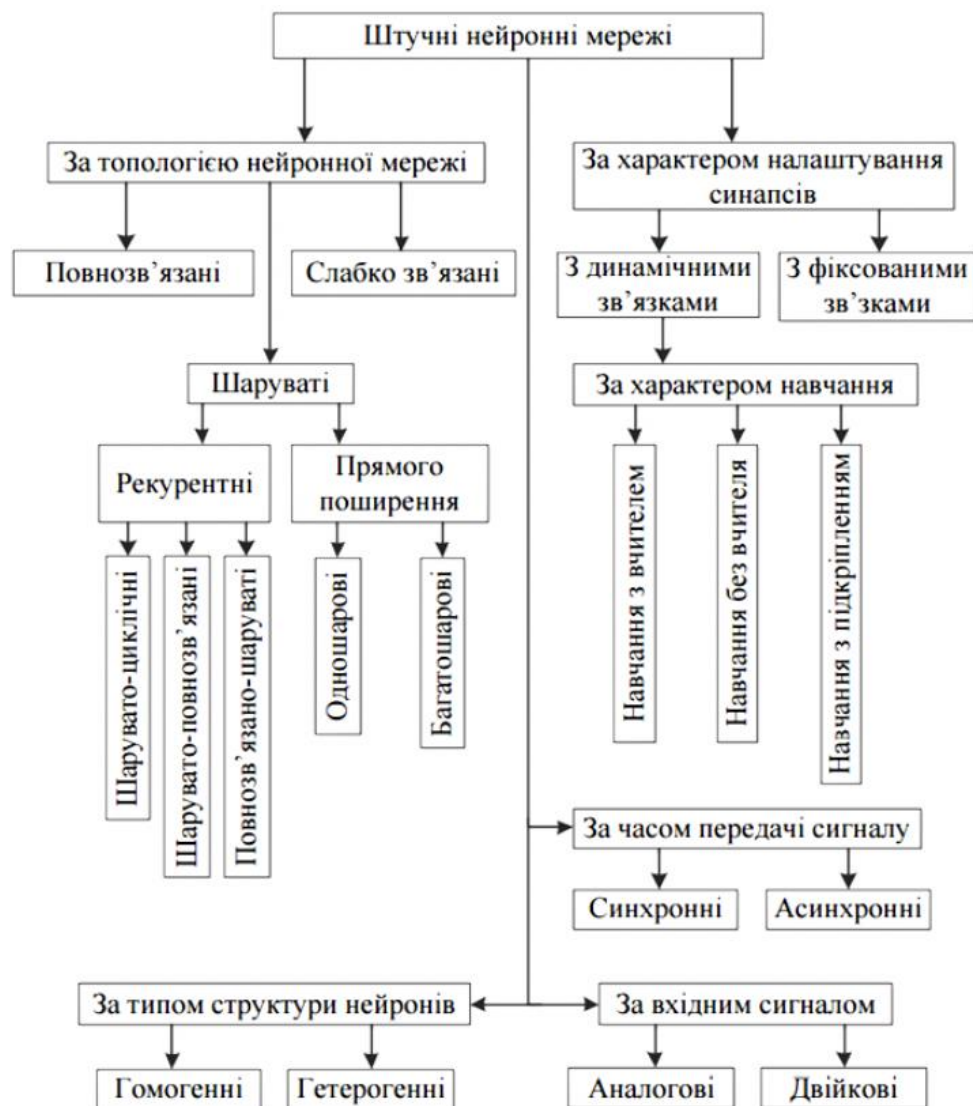


Рисунок 3 - Види штучних нейронних мереж

На противагу статичним архітектурам, існують динамічні архітектури штучних нейромереж, що реалізують рекурентну структуру з

використанням зворотніх зв'язків, завдяки чому стан мережі в кожний момент часу залежить від попереднього стану.

## **2.5 Принципи та методи навчання штучних нейронних мереж**

Головна перевага штучних нейронних мереж перед традиційними алгоритмами в тому, що вони не програмуються в звичайному розумінні цього слова, а навчаються. Навчання штучних нейронних мереж – це процес налаштування архітектури і ваг зв'язків між нейронами (параметрів) для виконання поставлених перед мережею завдань. Отже, в знаходженні коефіцієнтів зв'язків між нейронами полягає навчання. В данному процесі нейронна мережа виявляє залежності між вхідними даними й вихідними, робить узагальнення.

Зазвичай, передбачається наступна послідовність дій (під час навчання):

- В мережу надходять сигнали.
- Вільні параметри змінюються.
- Після того, як змінилася внутрішня структура нейрона мережа відповідає на ті ж самі вхідні сигнали вже іншим чином.

Алгоритмом навчання називають сукупність правила навчання нейронної мережі. Універсального алгоритму навчання, відповідного для всіх архітектур нейронних мереж, не існує.

Можна виділити два основні великі класи навчання: клас детерменірованих методів і клас стохастичних методів.

До класу детерменірованого навчання відносяться методи, в основі яких лежить ітеративна корекція зв'язків між нейронами мережі. Основним і найбільш затребуваним детерменірованим методом навчання мереж є метод зворотнього поширення помилки.

					ІАЛЦ.045490.004 ПЗ	Анк
						34
Изм.	Лис	№ докум.	Підпис	Дата		

До класу стохастичних методів відносяться ті, що змінюють параметри мережі випадковим чином. Але вони зберігають тільки ті зміни параметрів, які призвели до поліпшення результатів роботи. Алгоритми таких методів навчання здійснюються із застосуванням методу порівняння помилок. Для можливості застосування стохастичних методів повинна бути диференційована передавальна функція нейронів. Методи розглядаються як градієнтний спуск по поверхні помилки і є модифікацією класичного методу градієнтного спуску. В якості прикладу стохастичного методу можна привести такий алгоритм:

1. Обрати параметри мережі випадковим чином та підкорегувати їх на невелику випадкову величину. Обрахувати отримані значення виходів за запропонованими значеннями входів.
2. Порівняти ці виходи з бажаним результатом та обчислити різницю значень. Ця різниця називається помилкою. Ціль навчання поляє у тому, щоб мінімізувати значення помилки.
3. Значення помилки зменшено, корегування збережно. У негативному випадку значення корегування відкидається і вибирається заново.

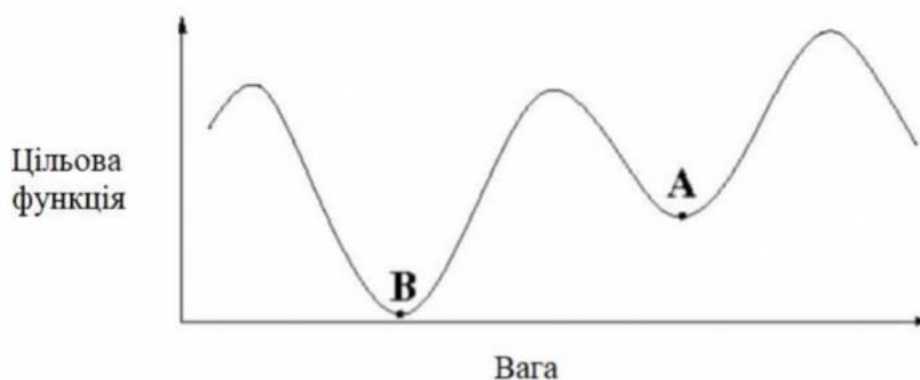


Рисунок 4 - Проблема локального мінімуму

Кроки 2 та 3 повторюють поки мережу не буде навчено.

Необхідно зауважити, що стохастичні методи навчання можуть потрапити до «пастки» локального мінімуму (Рис. 4).

Розглянемо зображення рисунку. Нехай початкове значення помилки рівно чи близьке до значення у точці А. Якщо випадкові кроки корегування дуже малі, то будь-які відхилення від точки А збільшать помилку та не будуть братися до уваги. Таким чином найменше значення у точці В ніколи не буде знайдено. При дуже великих випадкових кроках корегування параметрів мережі помилка буде піддаватися зміні так різко, що ніколи не закріпиться за одним із мінімумів.

Щоб уникати проблеми локального мінімуму можна поступово зменшити випадковий середній розмір кроків корегування. Коли середній розмір кроків завеликий, значення помилки буде приймати усі значення з рівною імовірністю. При плавному зменшенні розміру кроків буде дотримано умови при якому значення помилки буде на деякий час «застрягати» в точці В. Коли значення кроку буде ще більше зменшено, значення помилки буде «зупинятися» на короткий проміжок часу і в точці А і в точці В. При зменшенні кроку безперервно то у кінці буде досягнуто величину кроку, що буде достатньою для подолання локального мінімуму А, але не локального мінімуму В.

Математичне представлення навчання мереж зводиться до наступного вигляду:

Нехай ми маємо рівняння  $Y = f(X)$  яке вирішує задачу і набір вхідних-вихідних даних  $(X^1, Y^1), (X^2, Y^2) \dots (X^n, Y^n)$ . При навчанні мережі потрібно знайти таку функцію, що буде близькою до функції з помилкою. Тобто знаючи всі ці параметри завдання зводиться до задачі багатовимірної оптимізації функції

Існують три способи навчання: з вчителем; без вчителя, змішаний.

Детальніше розглянемо навчання мережі з учителем. У даному методі нейромережа має правильні відповіді (виходи мережі) на кожен вхідний приклад. Тобто, кожному прикладу з навчальної вибірки відповідає вектор. Він є характеристикою однозначної правильної відповіді, яка відразу подається на вихід мережі в обхід всієї її архітектури. Коли результат мережі отримано, алгоритм починає порівнювати результуючий вектор з правильною відповіддю. Потім здійснюється корекція подальших помилок. Ваги налаштовуються так, щоб мережа виробляла відповіді найбільш близькі до відомих правильних відповідей. Метод зворотного поширення помилки є класичним прикладом навчання з учителем і спирається саме на правило корекції помилки .

Алгоритм навчання з учителем:

1. Взяті дані для навчання мережі розділити на дві частини. Першою половиною буде навчальна вибірка, другою - тестова вибірка. (можливо розділити дані 70/30).
2. Прорахувати значення функції для навчальної вибірки та знайти значення функції помилки.
3. Провести корегування ваг синапсів у мережі.
4. Повторювати пункти 2 і 3 до тих пір поки значення функції помилки не буде мінімальним. Повторювати ці дії можна як для кожного екземпляра навчальної вибірки так і для всієї вибірки в цілому. У першому випадку навчання буде повільніше, але з більшою точністю. У другому випадку, показник точності занепадає, але навчання буде відбуватися швидше.
5. Перевірити всі значення за тестовою вибіркою, щоб зрозуміти наскільки добре система змогла узагальнити дані (навчитися).

					ІАЛЦ.045490.004 ПЗ	Анк
Изм.	Лис	№ докум.	Підпис	Дата		37



Необхідно відзначити, що при навчанні з учителем потрібна велика вибірка, щоб в достатній мірі сформувати робочу і гнучку нейронну мережу.

Типи навчання	Правило навчання	Архітектура	Алгоритм навчання	Задача
З вчителем	Корекція помилки	Одношаровий і багатшаровий перцептрон	Алгоритми навчання перцептрона Зворотне поширення	Класифікація образів Апроксимація функцій Передбачення, управління
	Больцман	Рекурентна	Алгоритм навчання Больцмана	Класифікація образів
	Хебб	Багатшарова прямого розповсюдження	Лінійний дискримінантний аналіз	Аналіз даних Класифікація образів
	Змагальні	Змагальня	Векторне квантування	Категоризація всередині класу Стиснення даних
		Мережа ART	ARTMap	Класифікація образів
Без вчителя	Корекція помилки	Багатшарова прямого розповсюдження	Проекція Саммона	Категоризація всередині класу Аналіз даних
	Хебб	Прямого поширення або змагання	Аналіз головних компонентів	Аналіз даних Стиснення даних
		Мережа Хопфілда	Навчання асоціативної пам'яті	Асоціативна пам'ять
	Змагальні	Змагальня	Векторне квантування	Категоризація Стиснення даних
		Мережа Кохонена	SOM Кохонена	Категоризація Аналіз даних
		Мережа ART	ART1, ART2	Категоризація
Змішані	Коррекция ошибки и соревнование	Мережа RBF	Алгоритм навчання RBF	Класифікація образів Апроксимація функцій Передбачення, управління

Таблиця 1 - Класифікація методів навчання нейромереж.

При навчанні без вчителя попередньо не існує правильних відповідей на кожен приклад навчальної вибірки. Мережа сама знаходить взаємозв'язки між елементами навчальної вибірки і здійснює автоматичне налаштування параметрів, яке призводить до появи однакових результатів її функціонування при досить близьких вхідних значеннях. Для коректування своїх ваг мережа не використовує зовнішніх впливів і внутрішньо контролює свою ефективність, шукаючи регулярність у

вхідних сигналах та відповідно до навчальної функції здійснює адаптацію. Вона має інформацію, яка закладена у топологію мережі та навчальні правила, відносно власної організації навіть без інформації про правильність чи неправильність дій. Конкуренція між нейронами формує основу для навчання. Правило Хеба і методи навчання при змаганні нейронів шляхом порівняння інтенсивності їх реакції є класичними прикладами методів навчання без учителя. При використанні змішаного навчання частина ваг визначається за допомогою самонавчання, у той час як інша визначається за допомогою навчання зі вчителем.

Ефективність навчання нейронної мережі залежить від кількох керованих факторів. Взагаліному розглядають три основні властивості, пов'язані з навчанням: ємність, обчислювальна і складність зразків.

- Ємність. Властивість вказує, які межі прийняття рішень можуть бути і скільки зразків може запам'ятати мережа.
- Обчислювальна складність. Властивість пов'язана з процесором ЕОМ (з його потужністю).
- Складність зразків. Властивість визначає кількість навчальних прикладів, необхідних для того, щоб досягнути здатність мережі до узагальнення.

Для навчання нейронних мереж є дуже багато правил, більшість з них є варіацією вже давно відомого правила навчання - правила Хеба.

Розглянемо правило Хебба. Опис цього правила був наведений в книзі Дональда Олдінга Хебба під назвою “Організація поведінки”.

Правило звучить наступним чином: коли нейрон отримує вхідний сигнал від іншого нейрону і вони обидва є високо активними (тобто мають такий самий знак), вагу між ними потрібно підсилити[1].

					ІАЛЦ.045490.004 ПЗ	Анк
						39
Изм.	Лис	№ докум.	Підпис	Дата		

Вага синаптичного з'єднання між одночасно збудженими двома нейронами з виходами ( $x_j$ ,  $y_i$ ) на  $k$ -тому кроці навчання, в іншому випадку - зменшується, згідно з формулою (2).

$$\Delta W_{ij}(k) = \eta x_j(k) y_i(k), \quad (2)$$

де  $\eta$  - коефіцієнт швидкості навчання.

Це правило застосовується при навчанні як "з вчителем" так і "без вчителя". Розглянемо правило Хопфілда. Воно є схожим до правила Хебба відмінність полягає в тому, що це правило визначає величину підсилення або послаблення. Правило звучить наступним чином: коли вихідний та вхідний сигнал нейрона одночасно є активними або неактивними, потрібно збільшити вагу з'єднання оцінкою навчання, в іншому випадку зменшуємо вагу оцінкою навчання[].

Розглянемо правило "дельта". Це правило є одним із найбільш використовуваним. Воно базується на ідеї неперервної зміни синаптичних ваг для зменшення "дельта" (різниці) між значенням бажаного та біжучого вихідного сигналу нейрона. Результатом використання правила "дельта" є мінімізація середньоквадратичної похибки мережі.

Похибка у цьому правилі, яка є отримана у вихідному прошарку перетворюється похідною переданої функції. Також вона послідовно поширюється назад на попередні прошарки для зміни синаптичних ваг. Триває цей процес зворотного поширення похибки мережі до етапу досягнення першого прошарку. Відома парадигма FeedForward BackPropagation успадкувала своє ім'я саме від методу обчислення похибок "дельта".

Невпорядкованість множини вхідних даних є важливим пунктом при використанні правила "дельта". При використанні добре впорядкованої або

					ІАЛЦ.045490.004 ПЗ	Анк
						40
Изм.	Лис	№ докум.	Підпис	Дата		

структурованої навчальної множини результат мережі може бути негативним, тобто не збігтися до бажаної точності і мережа буде вважатись нездатною до навчання.

Розглянемо правило градієнтного спуску. Це правило є подічним на правило, яке ми розглядали вище ("дельта"). Схожість визначається тим, що для змінювання похибки «дельта» перед тим, як застосувати її до ваг з'єднань використовують похідну від переданої функції. Правило градієнтного спуску є загально використовуване незважаючи на повільний процес навчання.

Розглянемо правило змагання. При навчанні цим методом вихідні нейрони змагаються між собою за активізацію. Це явище також називають, як правило "переможець отримує все". В біологічних нейронних мережах подібне навчання. Результатом навчання за допомогою змагання є кластеризовані вхідні дані.

Під час навчання відбуваються зміни синаптичних ваг нейрона-переможця. Ефект правила змагання здобувається за рахунок цієї зміни збереженого в мережі вектора синаптичних ваг нейрона-переможця. Переможцем вважається нейрон з найбільшим вихідним сигналом і він має можливість гальмувати своїх конкурентів також збуджувати сусідів. Під час використання вихідного сигналу нейрона-переможця тільки його сусідам та йому дозволяється коректувати свої ваги з'єднань.

Під час періоду навчання розмір області сусідства може змінюватись. Елемент-переможець визначається відповідно до найвищої відповідності до вхідного зразку, розподіл входів моделюють мережі Кохонена. Це правило зазвичай використовують в самоорганізованих картах.

Розглянемо метод зворотнього поширення помилки (Backprop). Це класичний метод навчання з учителем. Основна його ідея в тому, що

					ІАЛЦ.045490.004 ПЗ	Анк
						41
Изм.	Лис	№ докум.	Підпис	Дата		

сигнал помилки, після її обчислення на виході мережі, поширюється в зворотньому напрямку. Передавальна функція нейронів повинна бути диференційована. При цьому процесі налаштовуються синаптичні ваги з метою мінімізації помилки. Фактично, відбувається рух в багатовимірному просторі ваг до мінімуму помилки в сторону, протилежну градієнту. В процесі навчання рішення знаходяться циклічно.

Найбільш важливим неконтрольованим навчанням є кластеризація, яка створює різні кластери введення і зможе вносити будь-які нові дані до відповідного кластеру. Крім кластеризації є інші методи неконтрольованого навчання: виявлення аномалій, приховані змінні моделі, такі як алгоритм максимізації очікувань, метод моментів і методи поділу сліпих сигналів.

Метод навчання парцептрона заснований на корекції помилок, запропонований Френком Розенблатом ще у далекому 1957 році, являє собою такий метод навчання, при якому вага зв'язку змінюється, поки перцептрон видає правильний результат. Параметри впроваджують зміни у випадку неспівпадання вихідних значень з вхідними. Для обчислення величини корекції використовується різниця між реальним і бажаним значенням виходу мережі.

Дана модель використовує навчання з учителем, тобто навчальна множина складається із сукупності вхідних векторів для кожного з яких вказано вихідний вектор. Не зважаючи на деякі обмеження модель стала основою для багатьох сучасних більш складних алгоритмів навчання.

Метод конкурентного навчання полягає у змаганні кожного з вихідних нейронів мережі за активацію. З цього можна зробити висновок що з усіх вихідних нейронів до роботи приступає лише нейрон із самим великим виходом. Такий алгоритм має нагадування біологічних нейронних мереж. За допомогою методу конкуренції можна класифікувати вихідні

					ІАЛЦ.045490.004 ПЗ	Анк
Изм.	Лис	№ докум.	Підпис	Дата		42

дані де схожі приклади групуються до одного класу і подаються як один зразковий елемент. При цьому кожни нейрон із множини нейронів несе відповідаль лише за один клас. Загальне число класів з якими працює мережа дорівнюватиме число вихідних нейронів.

Конкурентне навчання дуже зручно використовуваних в задачах класифікації вхідних образів. У цьому випадку кожен нейрон вихідного шару відповідає за один образ.

## 2.6 Обґрунтування вибраного інструментарію

Запропонований метод розпізнавання реквізитів банківської платіжної картки реалізовано на мові програмування Python, що дозволила виконати об'єднання зазначених методів у комплекс при невеликому обсязі програмного коду. Python - це мова програмування високого рівня та загального призначення. Вона орієнтована на найбільш зручне читання коду на підвищення продуктивності розробника. Python має мінімалістичний синтаксис ядра, а його стандартна бібліотека включає багато корисних функцій.

Задача розпізнавання відноситься до класу задач машинного навчання. Вони ґрунтуються на результатах математичного аналізу, лінійної алгебри, методів оптимізації, математичної статистики. Мова Python має потужний математичний модуль math, бібліотеку складних обчислень scipy, бібліотеку лінійної алгебри numpy, а також кілька бібліотек машинного навчання, які містять всі основні алгоритми, інструменти для валідації алгоритмів та попередньої обробки даних. Крім того бібліотека pandas містить зручні засоби завантаження та представлення даних у вигляді data-фреймів, до яких можна застосовувати найрізноманітніші засоби обробки.

У даній роботі ще одним засобом обробки зображень є OpenCV. Це бібліотека обробки зображень, алгоритмів комп'ютерного зору, та

					ІАЛЦ.045490.004 ПЗ	Анк
						43
Изм.	Лис	№ докум.	Підпис	Дата		

чисельних алгоритмів загального призначення з відкритим вихідним кодом. Вона також реалізована на C++, проте розробляється і для Python. Для даної бібліотеки корисною функцією є можливість прямої роботи з веб-камерою та отримання з неї зображень, а також перетворення їх у різні формати.

Для створення користувацького інтерфейсу була обрана графічна бібліотека Tkinter. З її допомогою можна створювати програми з віконним інтерфейсом. Tkinter дозволяє створювати вікна, поля для вводу даних, програмне меню, елементи керування (кнопки, перемикачі), полотно для побудови графіків та виводу графічних об'єктів та ін.

Найперспективнішими методами, які використовуються для задачі розпізнавання символів, є нейронні мережі, метод найближчого сусіда, клітинні автомати та розпізнавання за шаблоном.

Ці методи допомагають вирішувати дуже багато практичних задач, зокрема, завдання розпізнавання. Наприклад, система, яка «читає» банківські чеки, по ефективності в кілька разів перевершує оператора.

Треба зважати на те, що нейронні мережі мають ряд недоліків: складність структури для проектування, застосування значних обчислювальних ресурсів. На відміну від них, метод найближчого сусіда зменшує кількість об'єктів навчальної вибірки та обчислювальні ресурси, що значно пришвидшує роботу програми.

Отже, для розпізнавання реквізитів банківських платіжних карток за їх зображеннями був обраний метод машинного навчання (з учителем). Задачу класифікації було вирішено за допомогою методу найближчих сусідів. За допомогою цього методу досягається потрібний результат.

					ІАЛЦ.045490.004 ПЗ	Анк
						44
Изм.	Лис	№ докум.	Підпис	Дата		

### 3 ОПИС РОЗРОБЛЕНОЇ ПРОГРАМИ

Розроблена програма та її функціонал задовільняють поставленим на початку розробки задачам. Використання системи можливо за допомогою мінімалістичного зручного інтерфейсу користувача. Нижче зупинимось на функціоналі та інтерфейсі докладніше.

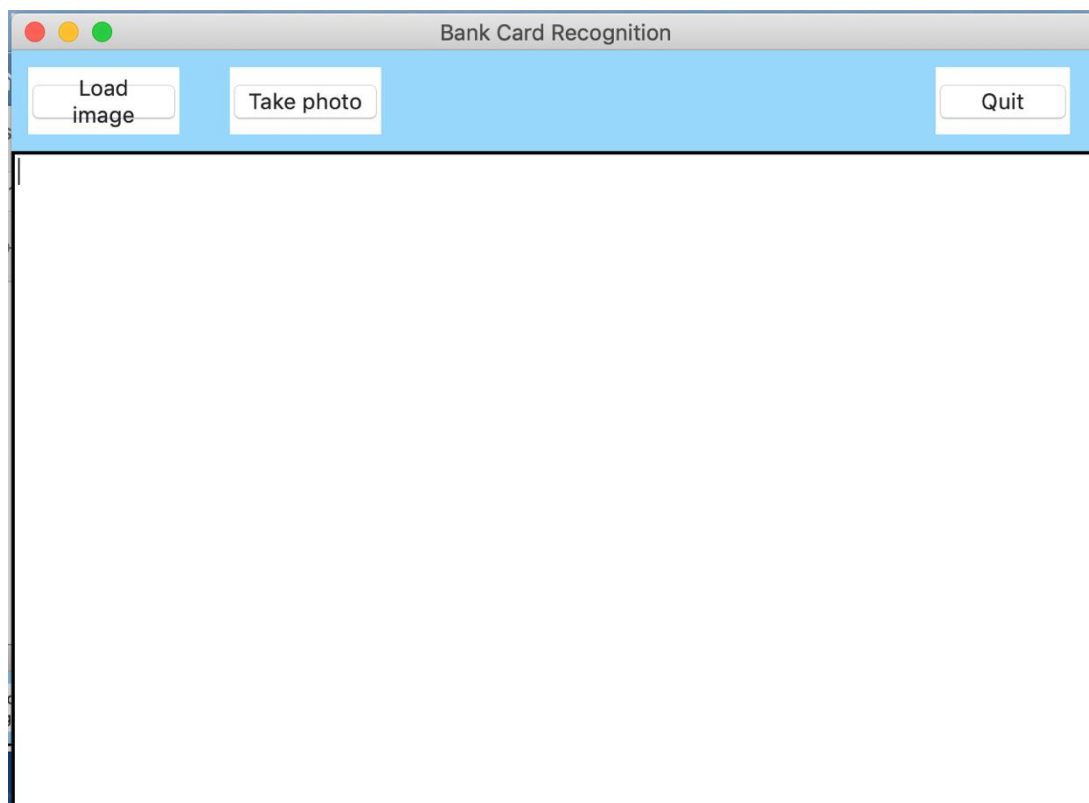


Рисунок 5 - Початковий стан програми

На старті програми відкривається головне вікно (Рис. 5), у якому відображені наступні кнопки:

- Завантаження зображення з файлової системи комп'ютера;
- Отримання моментального зображення за допомогою веб-камери;
- Коректний вихід з програми.

Після отримання зображення вікно програми має такий вигляд (Рис. 6):

					ІАЛЦ.045490.004 ПЗ	Анк
						45
Изм.	Лис	№ докум.	Підпис	Дата		





Рисунок 6 - Вікно програми після отримання зображення

Якщо на зображенні не виявлено банківської платіжної картки програма виведе повідомлення про некоректне зображення.

При розпізнаванні реквізитів банківської картки на екрані з'являється вікно, що відображає елементи, які були виявлені на зображенні (Рис.7). Елементи підсвічуються яскраво зеленим кольором.

Після розрахунку класифікаторів для виявлених елементів відбувається розпізнавання. Після розпізнавання отримані результати проходять важливий етап перевірки виявлених і розпізнаних реквізитів.

Перевірка на правильність розпізнавання такого реквізиту як номер банківської платіжної картки буде здійснюватися за допомогою алгоритму Луна. Перед використанням алгоритму здійснюється перевірка на кількість цифр(номер БПК складається із 16 десяткових цифр згрупованих по 4).

					ІАЛЦ.045490.004 ПЗ	Анк
						46
Изм.	Лис	№ докум.	Підпис	Дата		

Результатом роботи алгоритму розпізнавання терміну дії банківської платіжної картки повинен бути 4 десяткові цифри: дві - місяці, дві рік закінчення терміну дії. Вважається, що алгоритм видав правильну відповідь, якщо отримані 4 цифри збігаються з тими, що зображені на карті. Символ, що розділяє їх, не враховується і може бути будь-яким. Відмова від розпізнавання трактується як невірну відповідь.

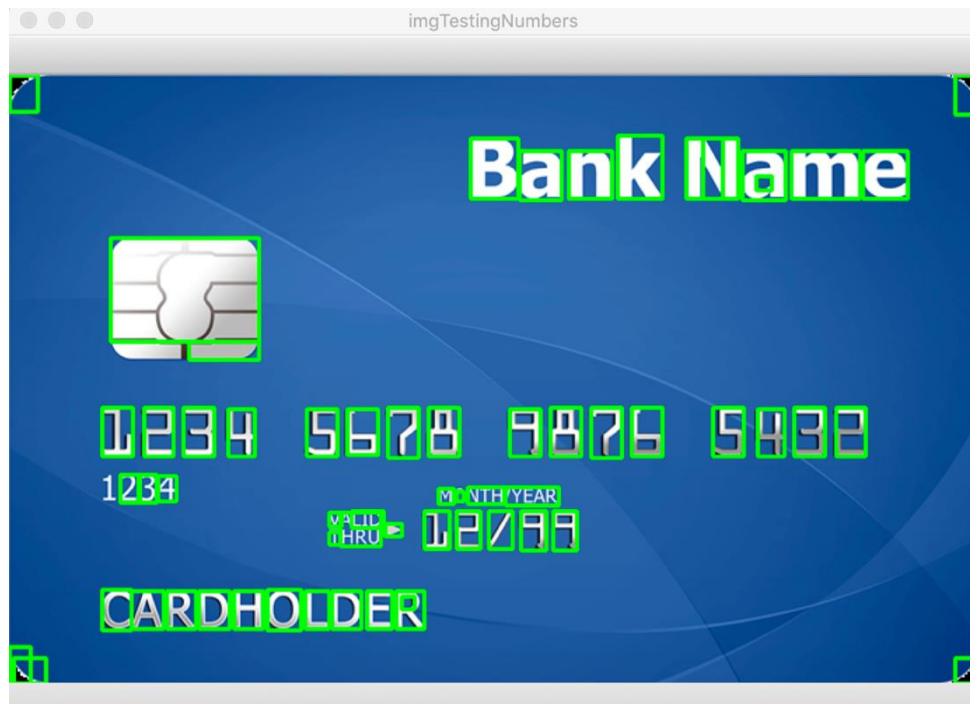


Рисунок 7 - Візуалізація елементів, які були виявлені на зображенні

Тренування моделі мережі з учителем відбувається наступним чином. Після запуску модуля навчання команда програма автоматично зафантажує підготовлений до початку тренування файл формату PNG. Цей фал містить зображення символів, які в майбутньому ми будемо розпізнавати, в “ідеальному середовищі”(символи з відповідним шрифтом на білому фоні). Програма розпізнає образ на зображенні формує класифікатор для цього образу і повідомляє про потребу ввести учителем символу з клавіатури, що відповідає виділеному образу. Після введення відповідного символу до всіх контурів, які було виявлено на зображенні

програма повідомляє про завершення тренування. Графічне зображення того, як відбувається навчання моделі мережі з учителем ви можете побачити на рисунку 8.



Рисунок 8 - Процес навчання модулі з учителем

Після завершення тренування програма виводить повідомлення про завершення тренування.

Після успішного розпізнавання реквізити банківської платіжної картки виводяться на екран (Рис. 9) в іншому випадку виводиться повідомлення про помилку.



Рисунок 9 - Вивід реквізитів банківської платіжної картки на екран

При натисканні кнопки виходу з програми закривається всі вікна програми для коректного завершення роботи.

					ІАЛЦ.045490.004 ПЗ	Анк
						49
Изм.	Лис	№ докум.	Підпис	Дата		

## ВИСНОВКИ

Результатом роботи над дипломним проектом на основі аналізу існуючих методів, алгоритмів та інструментів є розроблені алгоритм і програма розпізнавання реквізитів банківських платіжних карток на зображеннях. Під час виконання дипломного проекту було проведено огляд основних методів виявлення та розпізнавання символів, охарактеризовано і вирішено задачу розпізнавання реквізитів банківської картки. Для вирішення задачі і розробки програми були використані сучасні технології розробки програмного забезпечення, такі як: мова програмування Python, бібліотека комп'ютерного зору та машинного навчання OpenCV, бібліотека NumPy, графічний пакет Tkinter, та ін.

Вхідними даними програми є зображення у форматі PNG, що завантажується з файлової системи комп'ютера або може бути моментально отримане за допомогою веб-камери.

Використовуючи пакет Tkinter був розроблений зручний користувацький інтерфейс, що дозволяє легко керувати програмою.

Розроблена програма може бути використана як частина модуля проведення оплати онлайн. Також може бути основою для створення проекту для оптичного аналізу документів.

					ІАЛЦ.045490.004 ПЗ	Анк
						50
Изм.	Лис	№ докум.	Підпис	Дата		

## СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Методи\_розпізнавання\_тексту [Електронний ресурс] – Режим доступу до ресурсу: [https://uk.wikipedia.org/wiki/Методи\\_розпізнавання\\_тексту](https://uk.wikipedia.org/wiki/Методи_розпізнавання_тексту)
2. Кращі програми розпізнавання тексту [Електронний ресурс] – Режим доступу до ресурсу: <https://uk.soringgpcrepair.com/program-for-text-recognition/>
3. Что такое машинное обучение и в чем его особенность[Електронний ресурс] – Режим доступу до ресурсу: <https://www.rline.tv/svobodnyj-korrespondeent/obuchenie-mashinnoe/>
4. Моделирование. Тестирование, надёжность, контроль и диагностика компьютерных систем [Електронний ресурс]: учебное пособие для изучения дисциплин «Моделирование» и «Тестирование, надёжность, контроль и диагностика компьютерных систем» для иностранных студентов специальности «Компьютерные системы и сети», «Системное программирование» и «Специализированные компьютерные системы» / НТУУ «КПИ»; сост. В. В. Гроль, В. А. Романкевич, Е. Р. Потапова.– Електронні текстові дані (1 файл: 782.5 Кбайт). – Киев: НТУУ «КПИ», 2011
5. Гроль В.В., Романкевич В.О. Базові поняття і конструкції мови програмування Сі. Методичні вказівки до вивчення дисципліни “Моделювання” // Київ.– “Політехніка”, 2003.– 24с.
6. Самофалов К.Г., Романкевич А.М., Валуйский В.Н., Каневский Ю. С., Пиневи́ч М.М. Прикладная теория цифровых автоматов. – К.: Вища Школа. – 1987г. – 375с.
7. Гроль В.В., Романкевич В.А., Потапова Е.Р. Организация процедур логического моделирования цифровых блоков. Методические

					ІАЛЦ.045490.004 ПЗ	Анк
Изм.	Лис	№ докум.	Підпис	Дата		51

указания к изучению дисциплин «Моделирование», «Тестирование, надёжность, контроль и диагностика компьютерных систем» // Київ.– “Принт-центр”, 2007.– 44с.

8. Платёжная карта [Электронный ресурс] – Режим доступа до ресурсу: [https://ru.wikipedia.org/wiki/Платежная\\_карта](https://ru.wikipedia.org/wiki/Платежная_карта)
9. Як діє штучний інтелект і перспективи його розвитку[Електронний ресурс] – Режим доступу до ресурсу: <https://aiconference.com.ua/uk/news/printsipi-raboti-iskusstvennosgo-intellekta-i-perspektiva-ego-ispolzovaniya-92238>
10. Легко ли распознать информацию на банковской карточке? [Электронный ресурс] – Режим доступа до ресурсу: <https://habr.com/ru/company/smartengines/blog/2172607/>
11. Программный комплекс распознавания / К. Р.Ахметзянов, В. И. Сазонов, Ю. Н. Липин, А. А. Южаков. – С. 79–85.
12. Методы распознавания текстов [Электронный ресурс] – Режим доступа до ресурсу: <https://habr.com/ru/post/1124442/>